Foundation of Optimization
最优化原理

李晓鹏
Li Xiao-Peng

电子与信息工程学院
School of Electronic and Information Engineering

Semester–I 2024/25

第四章

1. 约束最优性条件

2. 惩罚函数法

引言

考虑一般的约束非线性规划问题

$$\begin{aligned} &\min\ f(\boldsymbol{x}) \\ &s.t.\ g_i(\boldsymbol{x}) \le 0, i=1,\cdots,m \cdots\cdots \text{(I)} \\ &\qquad h_j(\boldsymbol{x}) = 0, j=1,\cdots,p \end{aligned}$$

其中$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$.

约束优化问题中，自变量的取值受到限制，目标函数在无约束情况下的驻点很可能不在可行域内，**一般不能直接**用无约束优化的方法处理约束优化问题.

求解约束非线性规划问题的方法，即约束最优化方法，一般包括：惩罚函数法、可行方向法和二次逼近法.我们重点学习惩罚函数法.

目录

约束最优性条件

一、等式约束情形的求解方法及最优性条件

考虑如下模型:

$$\begin{aligned} \min \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & h_j(\boldsymbol{x}) = 0, \quad j = 1, \cdots, p \\ & \boldsymbol{x} \in \mathbb{R}^n \end{aligned}$$

其中 $f(\boldsymbol{x})$, $h_j(\boldsymbol{x})(j = 1, 2, \cdots, p)$ 具有一阶连续偏导数

高数中求解此类问题的方法: 引入 Lagrange 函数, 将上述问题转化为求解 Lagrange 函数的无约束极值问题.

约束最优性条件

针对上述问题，建立Lagrange函数：

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{v}) = f(\boldsymbol{x}) + v_1 h_1(\boldsymbol{x}) + \cdots + v_p h_p(\boldsymbol{x}),$$

其中$v_1, \cdots, v_p$ 为待定参数 Lagrange 乘子.

令 $\frac{\partial \mathcal{L}}{\partial x_i} = 0$ 得:

$$\nabla f(\boldsymbol{x}) + \sum_{j=1}^{p} v_j \nabla h_j(\boldsymbol{x}) = \mathbf{0},$$

再加上约束条件: $h_j(\boldsymbol{x}) = 0$, $j = 1, 2, \cdots, p$ 求解它们组成的方程组:

$$\begin{cases} \nabla f(\boldsymbol{x}) + \sum_{j=1}^{p} v_j \nabla h_j(\boldsymbol{x}) = \mathbf{0} \\ h_j(\boldsymbol{x}) = 0, \quad j = 1, 2, \cdots, p \end{cases}$$

得最优解 $\boldsymbol{x}^*$ 及最优乘子 $\boldsymbol{v}^*$.

约束最优性条件

Lagrange乘子法(解析法)步骤

1) 作Lagrange函数:

$$\mathcal{L}(\boldsymbol{x}) = f(\boldsymbol{x}) + \sum_{j=1}^{p} v_j h_j(\boldsymbol{x}),$$

其中 $\boldsymbol{v} = (v_1, v_2, \cdots, v_p)^T$ 称为Lagrange乘子向量

2) 令 $\frac{\partial \mathcal{L}}{\partial x} = 0$, 得

$$\nabla f(\boldsymbol{x}) + \sum_{j=1}^{p} v_j \nabla h_j(\boldsymbol{x}) = \mathbf{0}$$

3) 添加上原问题中约束条件: $h_j(\boldsymbol{x}) = 0, j = 1, 2, \cdots, p$
4) 求解由上述 $n + p$ 个方程构成的方程组, 得极值点 $\boldsymbol{x}^*$ 及最优的Lagrange乘子 $\boldsymbol{v}^* = (v_1^*, v_2^*, \cdots, v_p^*)^T$

约束最优性条件

定理 4.1：求解等式约束问题最优性条件

设$\boldsymbol{x}^*$是等式约束问题的局部最优解; $f(\boldsymbol{x})$, $h_j(\boldsymbol{x})(j=1,2,\cdots p)$在$\boldsymbol{x}^*$ 的某邻域内一阶连续可微, 且$\nabla h_j\left(\boldsymbol{x}^*\right)(j=1,2,\cdots p)$线性无关, 则存在一组实数$v_1^*, v_2^*, \cdots v_p^*$使得

$$\nabla f\left(\boldsymbol{x}^*\right)+\sum_{j=1}^p v_j^* \nabla h_j\left(\boldsymbol{x}^*\right)=\mathbf{0}. \boxed{\text{一阶必要条件}}$$

约束最优性条件

例 4.1: 用Lagrange乘子法求解下列问题:

$$\begin{aligned} \min\ & f(\boldsymbol{x}) = x_1 - 2x_2 + 3x_3 \\ \text{s.t.}\ & x_1^2 + x_2^2 + x_3^2 = 1 \end{aligned}$$

例 4.1: 用Lagrange乘子法求解下列问题:

$$\begin{aligned} \min\ & f(\boldsymbol{x}) = x_1 - 2x_2 + 3x_3 \\ \text{s.t.}\ & x_1^2 + x_2^2 + x_3^2 = 1 \end{aligned}$$

解: 作Lagrange函数
$\mathcal{L}(\boldsymbol{x}, v) = x_1 - 2x_2 + 3x_3 + v\left(x_1^2 + x_2^2 + x_3^2 - 1\right)$, 其中 v 为待定Lagrange乘子.
令 $\nabla_{\boldsymbol{x}}(\boldsymbol{x}, v) = \mathbf{0}$ 得:

$$\frac{\partial \mathcal{L}}{\partial x_1} = 1 + 2vx_1 = 0 \quad \cdots\cdots(1)$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = -2 + 2vx_2 = 0 \quad \cdots\cdots(2)$$

$$\frac{\partial \mathcal{L}}{\partial x_3} = 3 + 2vx_3 = 0 \quad \cdots\cdots(3)$$

约束最优性条件

再加上约束条件:

$$x_1^2 + x_2^2 + x_3^2 = 1 \quad \cdots\cdots (4)$$

解(1)–(4)组成的方程组得: $\boldsymbol{x}_1 = \left(\frac{1}{\sqrt{14}}, -\frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}}\right)^T, v_1 = -\frac{\sqrt{14}}{2}$,
对应的目标函数值为: $\sqrt{14}$
或 $\boldsymbol{x}_2 = \left(-\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, -\frac{3}{\sqrt{14}}\right)^T, v_2 = \frac{\sqrt{14}}{2}$, 对应的目标函数值为: $-\sqrt{14}$

故原问题的最优解为:

$$\boldsymbol{x}^* = \boldsymbol{x}^2 = \left(-\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, -\frac{3}{\sqrt{14}}\right)^T$$

二、一般约束情形的最优性条件

考虑约束非线性规划问题

$$\begin{aligned} \min\ & f(\boldsymbol{x}) \\ \text{s.t. } & g_i(\boldsymbol{x}) \leq 0, i = 1, \cdots, m \\ & h_j(\boldsymbol{x}) = 0, j = 1, \cdots, p \cdots\cdots (I) \\ & \boldsymbol{x} \in \mathbb{R}^n \end{aligned}$$

其中 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$, 用 $\boldsymbol{\Omega}$ 表示可行域, 则

$$\boldsymbol{\Omega} = \{\boldsymbol{x} \in \mathbb{R}^n \,|g_i(\boldsymbol{x}) \leq 0, h_j(\boldsymbol{x}) = 0, i = 1, \cdots, m, j = 1, \cdots, p\}$$

定义相应的 Lagrange 函数:

$$\mathcal{L}(\boldsymbol{x}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} u_i g_i(\boldsymbol{x}) + \sum_{j=1}^{p} v_j h_j(\boldsymbol{x})$$

约束最优性条件

有效约束:
若(I)中的一个可行点$\boldsymbol{x}^0$使得某个不等式约束$g_i(\boldsymbol{x}) \le 0$变成等式, 即 $g_i\left(\boldsymbol{x}^0\right) = 0$, 则$g_i(\boldsymbol{x}) \le 0$称为$\boldsymbol{x}^0$的有效约束, 也称紧约束条件.

有效约束指标集:

$$\boldsymbol{I} = \boldsymbol{I}\left(\boldsymbol{x}^0\right) = \left\{i \,\middle|\, g_i\left(\boldsymbol{x}^0\right) = 0, i = 1, 2, \cdots, m\right\}$$

无效约束:
若(I)中的一个可行点$\boldsymbol{x}^0$使得某个不等式约束$g_k(\boldsymbol{x}^0) < 0$, 则$g_k(\boldsymbol{x}) \le 0$称为点$\boldsymbol{x}^0$处的无效约束, 也称松约束条件.

约束最优性条件

例 4.2: 考察问题

$$
\begin{aligned}
\min\ & f(\boldsymbol{x}) = (x_1-2)^2 + x_2^2 \\
\text{s.t.}\ & g_1(\boldsymbol{x}) = -x_1 \le 0 \\
& g_2(\boldsymbol{x}) = -x_2 \le 0 \\
& g_3(\boldsymbol{x}) = x_1^2 + x_2^2 - 1 \le 0
\end{aligned}
$$

设 $\boldsymbol{x}^0 = (1,0)^T$, 判断上述哪些条件是有效约束?

约束最优性条件

例 4.2: 考察问题

$$
\begin{aligned}
\min\ f(\boldsymbol{x}) &= (x_1-2)^2 + x_2^2 \\
\text{s.t. } g_1(\boldsymbol{x}) &= -x_1 \le 0 \\
g_2(\boldsymbol{x}) &= -x_2 \le 0 \\
g_3(\boldsymbol{x}) &= x_1^2 + x_2^2 - 1 \le 0
\end{aligned}
$$

设 $\boldsymbol{x}^0 = (1,0)^T$, 判断上述哪些条件是有效约束?

解: 将 $\boldsymbol{x}^0$ 代入上述约束条件, 有 $g_1(\boldsymbol{x}^0) < 0$, $g_2(\boldsymbol{x}^0) = 0$, $g_3(\boldsymbol{x}^0) = 0$

所以 $g_2(\boldsymbol{x}) \le 0$, $g_3(\boldsymbol{x}) \le 0$ 为该点的有效约束条件.

有效约束指标集为: $\boldsymbol{I} = \{2, 3\}$.

约束最优性条件

定理 4.2：(一阶必要条件)
设$\boldsymbol{x}^*$为问题(I)的局部最优解，$\boldsymbol{I}^* = \{i | g_i(\boldsymbol{x}^*) = 0, 1 \le i \le m\}$；$f(\boldsymbol{x})$，$g_i(\boldsymbol{x})$，$h_j(\boldsymbol{x})$在$\boldsymbol{x}^*$点可微，且$\nabla h_j(\boldsymbol{x}^*)\ (j = 1, \cdots, p)$，$\nabla g_i(\boldsymbol{x}^*)$，$i \in \boldsymbol{I}^*$线性无关，则存在最优乘子$u_i^*(i = 1, \cdots, m)$，$v_j^*(j = 1, \cdots, p)$，使得：

(1) $\nabla f(\boldsymbol{x}^*) + \sum_{i=1}^m u_i^* \nabla g_i(\boldsymbol{x}^*) + \sum_{j=1}^p v_j^* \nabla h_j(\boldsymbol{x}^*) = \boldsymbol{0}$

(2) $u_i^* g_i(\boldsymbol{x}^*) = 0, i = 1, 2, \cdots, m$

(3) $u_i^* \ge 0, \quad i = 1, 2, \cdots, m$

【注】上述条件是由Kuhn和Tuchker于1951年提出，故一阶必要条件称为K-T条件，满足上述条件的点称为K-T点.1939年，Karush也提出了类似的约束最优性条件，故一阶必要条件也称作K-K-T条件，K-T点也称为K-K-T点.

约束最优性条件

一阶必要条件的说明:

(1) 满足K-T条件的可行点称为K-T点;

(2) 无效约束条件对应的Lagrange乘子为0; 因此, K-T条件便简化为:
存在Lagrange乘子$u_i^* \geq 0 \, (i \in \boldsymbol{I}^*)$和$v_j^* \, (j = 1, \cdots, p)$使得

$$\nabla f\left(\boldsymbol{x}^*\right) + \sum_{i \in \boldsymbol{I}^*} u_i^* \nabla g_i\left(\boldsymbol{x}^*\right) + \sum_{j=1}^{p} v_j^* \nabla h_j(\boldsymbol{x}^*) = \mathbf{0}$$

(3) 整体有效约束函数的梯度线性无关称为Kuhn–Tucker约束规范.

【注】该定理说明, 若最优点处对应的约束规范成立, 则它一定是K-T点; 若约束规范不满足, 则最优点不一定是K-T点.

约束最优性条件

判断某可行点是否是K-T点的步骤:

1) 找出不等式约束中的有效约束条件;
2) 求出K-T条件中相关梯度向量:

$$\begin{aligned}&\nabla f(\boldsymbol{x}^k),\\&\nabla g_i(\boldsymbol{x}^k), i \in \boldsymbol{I}(\boldsymbol{x}^k)\\&\nabla h_j(\boldsymbol{x}^k), j = 1, 2, \cdots, p\end{aligned}$$

3) 代入K-T条件的简化形式, 求出Lagrange乘子

$$u_i^*\,(i \in \boldsymbol{I}^*)\,,\;\; v_j^*\,(j = 1, \cdots, p)\,;$$

4) 若不等式约束中的有效约束条件对应的Lagrange乘子均非负, 那么该点是K-T点; 否则不是K-T点.

例 4.3: 考察问题

$$
\begin{aligned}
\min\ & f(\boldsymbol{x}) = (x_1-3)^2 + (x_2-2)^2 \\
s.t.\ & g_1(\boldsymbol{x}) = x_1^2 + x_2^2 - 5 \le 0 \\
& g_2(\boldsymbol{x}) = -x_1 \le 0 \\
& g_3(\boldsymbol{x}) = -x_2 \le 0 \\
& h_1(\boldsymbol{x}) = x_1 + 2x_2 - 4 = 0
\end{aligned}
$$

判断可行点$\boldsymbol{x}^0 = (2,1)^T$是否是该问题的K-T点.

约束最优性条件

例 4.3: 考察问题

$$
\begin{aligned}
\min\ f(\boldsymbol{x}) &= (x_1-3)^2+(x_2-2)^2 \\
s.t.\ g_1(\boldsymbol{x}) &= x_1^2+x_2^2-5 \le 0 \\
g_2(\boldsymbol{x}) &= -x_1 \le 0 \\
g_3(\boldsymbol{x}) &= -x_2 \le 0 \\
h_1(\boldsymbol{x}) &= x_1+2x_2-4=0
\end{aligned}
$$

判断可行点$\boldsymbol{x}^0=(2,1)^T$是否是该问题的K-T点.

解:
(1)计算$\boldsymbol{x}^0$点的有效约束：易验证对$\boldsymbol{x}^0$而言, 只有$\boldsymbol{g}_1(x)\le 0$是有效约束.

(2) 计算$\boldsymbol{x}^0$点的相关函数的梯度：

$$\nabla f(\boldsymbol{x}^0) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}, \nabla g_1(\boldsymbol{x}^0) = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \nabla h_1(\boldsymbol{x}^0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

(3) 求Lagrange乘子：按照K-T条件应有

$$\begin{pmatrix} -2 \\ -2 \end{pmatrix} + u_1 \begin{pmatrix} 4 \\ 2 \end{pmatrix} + v_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

解之得：$u_1 = \frac{1}{3}, \quad v_1 = \frac{2}{3}$.又因$u_1 = \frac{1}{3} > 0$, 所以K-T条件成立.

综上可知, $\boldsymbol{x}^0 = (2, 1)^T$是一个K-T点.

约束最优性条件

例 4.4: 考察问题

$$
\begin{aligned}
\min\ f(\boldsymbol{x}) &= (x_1-3)^2+(x_2-2)^2 \\
s.t.\ g_1(\boldsymbol{x}) &= x_1^2+x_2^2-5 \le 0 \\
g_2(\boldsymbol{x}) &= -x_1 \le 0 \\
g_3(\boldsymbol{x}) &= -x_2 \le 0 \\
h_1(\boldsymbol{x}) &= x_1+2x_2-4 = 0
\end{aligned}
$$

判断可行点$\boldsymbol{x}^1=(0,2)^T$是否是该问题的K-T点.

约束最优性条件

例 4.4: 考察问题

$$\begin{aligned}
\min\ & f(\boldsymbol{x}) = (x_1-3)^2+(x_2-2)^2 \\
s.t.\ & g_1(\boldsymbol{x}) = x_1^2+x_2^2-5 \le 0 \\
& g_2(\boldsymbol{x}) = -x_1 \le 0 \\
& g_3(\boldsymbol{x}) = -x_2 \le 0 \\
& h_1(\boldsymbol{x}) = x_1+2x_2-4 = 0
\end{aligned}$$

判断可行点$\boldsymbol{x}^{\mathrm{l}}=(0,2)^T$是否是该问题的K-T点.

解:

(1)计算$\boldsymbol{x}^1$点的有效约束：易验证对$\boldsymbol{x}^1$而言, 只有$g_2(\boldsymbol{x})\le 0$是有效约束.

约束最优性条件

(2)计算$\boldsymbol{x}^1$点的相关函数的梯度：

$$\nabla f(\boldsymbol{x}^1) = \begin{pmatrix} -6 \\ 0 \end{pmatrix}, \nabla g_2(\boldsymbol{x}^1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \nabla h_1(\boldsymbol{x}^1) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

(3)求Lagrange乘子：按照K-T条件应有

$$\begin{pmatrix} -6 \\ 0 \end{pmatrix} + u_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + v_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

解之得：$u_2 = -6$, $v_1 = 0$.但因$u_2 = -6 < 0$, 所以K-T条件不成立.

综上可知, $\boldsymbol{x}^1 = (0, 2)^T$不是K-T点.

约束最优性条件

例 4.5: 考察问题

$$\begin{aligned} \min\ & f(\boldsymbol{x}) = -x_1 \\ s.t\ & g_1(\boldsymbol{x}) = x_1^3 - x_2 \le 0 \\ & g_2(\boldsymbol{x}) = x_2 \le 0 \end{aligned}$$

判断$\boldsymbol{x}^* = (0,0)^T$处K-T条件是否成立?

例 4.5: 考察问题

$$\begin{aligned} \min\ & f(\boldsymbol{x}) = -x_1 \\ s.t\ & g_1(\boldsymbol{x}) = x_1^3 - x_2 \le 0 \\ & g_2(\boldsymbol{x}) = x_2 \le 0 \end{aligned}$$

判断$\boldsymbol{x}^* = (0,0)^T$处K-T条件是否成立?

解: 易验证对$\boldsymbol{x}^*$而言, $g_1(\boldsymbol{x}) \le 0$, $g_2(\boldsymbol{x}) \le 0$是紧约束.

注意到$\nabla f(\boldsymbol{x}^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \nabla g_1(\boldsymbol{x}^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \nabla g_2(\boldsymbol{x}^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. 因此, 对任意$u_1$, u_2:

$$\nabla f(\boldsymbol{x}^*) + u_1 \nabla g_1(\boldsymbol{x}^*) + u_2 \nabla g_2(\boldsymbol{x}^*) = \begin{pmatrix} -1 \\ -u_1+u_2 \end{pmatrix} \ne \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

所以$\boldsymbol{x}^*$处K-T条件不成立, 即该点不是K-T点.

定理 4.3：凸规划问题的全局最优解判定条件

对于凸规划

$$\begin{aligned} &\min\ f(\boldsymbol{x}) \\ &s.t.\ g_i(\boldsymbol{x}) \le 0, i = 1, \cdots, m \\ &\qquad h_j(\boldsymbol{x}) = 0, j = 1, \cdots, p \end{aligned}$$

设$f(\boldsymbol{x})$, $g_i(\boldsymbol{x})(i = 1, 2, \cdots, m)$是可微的凸函数，$h_j(\boldsymbol{x})(j = 1, 2, \cdots, p)$ 为线性函数，且$\boldsymbol{x}^* \in \boldsymbol{S}$. 如果$\boldsymbol{x}^*$是K-T点，则$\boldsymbol{x}^*$是全局最优点.

约束最优性条件

例 4.6: 考察问题

$$
\begin{aligned}
\min\ & f(\boldsymbol{x}) = (x_1-3)^2 + (x_2-2)^2 \\
\text{s.t. } & g_1(\boldsymbol{x}) = x_1^2 + x_2^2 - 5 \le 0 \\
& g_2(\boldsymbol{x}) = -x_1 \le 0 \\
& g_3(\boldsymbol{x}) = -x_2 \le 0 \\
& h_1(\boldsymbol{x}) = x_1 + 2x_2 - 4 = 0
\end{aligned}
$$

判断可行点 $\boldsymbol{x}^0 = (2,1)^T \in \boldsymbol{S}$ 是否是该问题的K-T点.

约束最优性条件

例 4.6: 考察问题

$$
\begin{aligned}
\min\ & f(\boldsymbol{x}) = (x_1-3)^2 + (x_2-2)^2 \\
\text{s.t. } & g_1(\boldsymbol{x}) = x_1^2 + x_2^2 - 5 \le 0 \\
& g_2(\boldsymbol{x}) = -x_1 \le 0 \\
& g_3(\boldsymbol{x}) = -x_2 \le 0 \\
& h_1(\boldsymbol{x}) = x_1 + 2x_2 - 4 = 0
\end{aligned}
$$

判断可行点 $\boldsymbol{x}^0 = (2,1)^T \in \boldsymbol{S}$ 是否是该问题的K-T点.

解： 由例4.4知点$\boldsymbol{x}^0 = (2,1)^T$是K-T点.又因函数$f(\boldsymbol{x})$, $g_i(\boldsymbol{x})$, $h_1(\boldsymbol{x})$都满足定理4.3中的条件. 故该问题是凸规划问题, 故K-T点也是最优解, 所以$\boldsymbol{x}^0 = (2,1)^T$是全局最优解.

目录

惩罚函数法–外点法

基本思想:
通过引入惩罚函数, 将所求解的约束非线性规划问题转化为一系列无约束非线性规划问题.

根据约束的特点(等式或不等式), 构造某种惩罚函数, 然后把它加到目标函数中去, 将约束问题的求解化为一系列无约束问题, 求解这一系列无约束优化问题, 可得到约束非线性优化问题的解.准确地说, 是将这些无约束问题的极小点依次作为迭代点.

将优化问题转化为一系列无约束优化来求解约束优化的技术称为序列无约束极小化技术(sequentail unconstrained minimization technique, SUMT).

惩罚函数法–外点法

辅助函数:

$$F(\boldsymbol{x},\sigma)=f(\boldsymbol{x})+\sigma P(\boldsymbol{x}).$$

上述惩罚策略, 对于在一系列极小化问题$\min_{\boldsymbol{x}\in\mathbb{R}^n} F(\boldsymbol{x},\sigma)$的求解中企图违反约束的极小点给予很大的目标函数值, 迫使无约束的极小点逐渐靠近可行域, 或者一直保持在可行域内, 直到收敛到原来约束优化的极小点或最优解.

根据惩罚函数表达式(构造方法的不同), 形成不同的罚函数法, 包括:
外点(惩)罚函数法, 内点(惩)罚函数法, 以及混合(惩)罚函数法.

惩罚函数法–外点法

一、外点惩罚函数法一外点法

考虑如下问题：

$$
\begin{aligned}
\min\ & f(\boldsymbol{x}) \\
s.t.\ & g_i(\boldsymbol{x}) \leq 0, i = 1, \cdots, m \\
& h_j(\boldsymbol{x}) = 0, j = 1, \cdots, p
\end{aligned}
$$

构造辅助函数：$F(\boldsymbol{x}, \sigma) = f(\boldsymbol{x}) + \sigma P(\boldsymbol{x})$，其中$\sigma > 0$是常数，称为惩罚因子；$\sigma P(\boldsymbol{x})$称为惩罚项；$P(\boldsymbol{x})$是定义在$\mathbb{R}^n$上的函数，称为惩罚函数，并且满足：

1. $P(\boldsymbol{x})$是连续的;
2. 对任意的$\boldsymbol{x} \in \mathbb{R}^n$，都有$P(\boldsymbol{x}) \geq 0$;
3. 当且仅当$\boldsymbol{x}$是可行点时，$P(\boldsymbol{x}) = 0$.

惩罚函数法–外点法

$$P(\boldsymbol{x}) = \sum_{i=1}^{m}\left[\max\{0, g_i(\boldsymbol{x})\}\right]^2 + \sum_{j=1}^{p}\left|h_j(\boldsymbol{x})\right|^2$$

分析:

- 当$\boldsymbol{x} \in \boldsymbol{S}$时, $P(\boldsymbol{x}) = 0$, $F(\boldsymbol{x}, \sigma) = f(\boldsymbol{x})$.
- 当$\boldsymbol{x} \notin \boldsymbol{S}$时, $P(\boldsymbol{x}) > 0$, 又因σ是极大的正数.因此, $\boldsymbol{x}$很难成为$F(\boldsymbol{x}, \sigma)$ 的极小点.

因此, 按上策略迫使$F(\boldsymbol{x}, \sigma)$的极小点逐渐靠近可行域, 一旦成为可行点, 即是最优解.具体而言, 对违反约束的点(即不可行点)在目标函数中加入相应的惩罚, 可行点不予惩罚, 这种方法的迭代点一般在约束优化的可行域外部移动, 故称此类算法为外点法.

惩罚函数法–外点法

定理 4.4:

若$\boldsymbol{x}(\sigma)$是$\min\limits_{\boldsymbol{x}\in\mathbb{R}^n} F(\boldsymbol{x},\sigma)$的最优解, 且$\boldsymbol{x}(\sigma)\in\boldsymbol{S}$, 则$\boldsymbol{x}(\sigma)$也是$\min\limits_{\boldsymbol{x}\in\boldsymbol{S}} f(\boldsymbol{x})$的最优解.

惩罚函数法–外点法

定理 4.4：
若$\boldsymbol{x}(\sigma)$是$\min\limits_{\boldsymbol{x}\in\mathbb{R}^n} F(\boldsymbol{x},\sigma)$的最优解，且$\boldsymbol{x}(\sigma)\in\boldsymbol{S}$，则$\boldsymbol{x}(\sigma)$也是$\min\limits_{\boldsymbol{x}\in\boldsymbol{S}} f(\boldsymbol{x})$的最优解.

证： $\forall\boldsymbol{x}\in\boldsymbol{S}$，因为$\boldsymbol{x}(\sigma)$是$\min_{\boldsymbol{x}\in\mathbb{R}^n} F(\boldsymbol{x},\sigma)$的最优解，所以

$$F(\boldsymbol{x}(\sigma),\sigma)\le F(\boldsymbol{x},\sigma),\ \ \forall\boldsymbol{x}\in\boldsymbol{S}.$$

注意到对于$\forall\boldsymbol{x}\in\boldsymbol{S}$，有$P(\boldsymbol{x}(\sigma))=0$.因此，对于$\forall\boldsymbol{x}\in\boldsymbol{S}$：

$$\begin{aligned}
f(\boldsymbol{x}(\sigma)) &= f(\boldsymbol{x}(\sigma))+\sigma P(\boldsymbol{x}(\sigma))\\
&= F(\boldsymbol{x}(\sigma),\sigma)\le F(\boldsymbol{x},\sigma),\ \ \forall\boldsymbol{x}\in\boldsymbol{S}\\
&= f(\boldsymbol{x}(\sigma))+\sigma P(\boldsymbol{x}(\sigma))=f(\boldsymbol{x}).
\end{aligned}$$

惩罚函数法–外点法

2. 解析法

$$\begin{aligned} \min\ & f(\boldsymbol{x}) \\ s.t.\ & g_i(\boldsymbol{x}) \le 0, \quad i = 1, \cdots, m \\ & h_j(\boldsymbol{x}) = 0, \quad j = 1, \cdots, p \end{aligned}$$

(1) 构造辅助函数：$F(\boldsymbol{x}, \sigma) = f(\boldsymbol{x}) + \sigma P(\boldsymbol{x})$, 其中$\sigma > 0$是很大的正数，$P(\boldsymbol{x}) = \sum_{i=1}^{m} \left[\max\left\{0, g_i(\boldsymbol{x})\right\}\right]^2 + \sum_{j=1}^{p} \left|h_j(\boldsymbol{x})\right|^2$.

(2) 求解：$\min_{\boldsymbol{x} \in \mathbb{R}^n} F(\boldsymbol{x}, \sigma) = f(\boldsymbol{x}) + \sigma P(\boldsymbol{x})$得最优解$\boldsymbol{x}(\sigma)$.

(3) 令$\sigma \to +\infty$, 有$\boldsymbol{x}(\sigma) \to \boldsymbol{x}^*$, 即得原问题的最优解.

惩罚函数法–外点法

例 4.7: 求解等式约束问题:

$$\begin{aligned} \min\ & f(\boldsymbol{x}) = x_1^2 + x_2^2 \\ \text{s.t. } & x_1 + x_2 - 2 = 0 \end{aligned}$$

惩罚函数法–外点法

例 4.7: 求解等式约束问题:

$$\min \ f(\boldsymbol{x}) = x_1^2 + x_2^2$$
$$\text{s.t. } x_1 + x_2 - 2 = 0$$

解: 构造罚函数和辅助函数:

$$F(\boldsymbol{x}, \sigma) = x_1^2 + x_2^2 + \sigma(x_1 + x_2 - 2)^2$$

其中 σ 是很大的正数.
令:

$$\frac{\partial F}{\partial x_1} = \frac{\partial F}{\partial x_2} = 0$$

得: $x_1(\sigma) = x_2(\sigma) = \frac{2\sigma}{2\sigma+1}$.

又因该点处 $\frac{\partial^2 F}{\partial x_1^2} = 2\left(\sigma + 1\right)$, $\frac{\partial^2 F}{\partial x_2^2} = 2\left(\sigma + 1\right)$, $\frac{\partial^2 F}{\partial x_1 \partial x_2} = 2\sigma$

并且 $\sigma > 0$, 可见 $\nabla^2 F = \begin{pmatrix} 2(\sigma+1) & 2\sigma \\ 2\sigma & 2(\sigma+1) \end{pmatrix}$ 正定.

因此 $F(\boldsymbol{x}, \sigma)$ 极小值点为 $x(\sigma) = \left(\frac{2\sigma}{2\sigma+1}, \frac{2\sigma}{2\sigma+1}\right)^T$

令 $\sigma \to +\infty$ 时, 有:

$$(x_1(\sigma), x_2(\sigma))^T \to (x_1^*, x_2^*)^T = (1, 1)^T$$

即: 无约束问题的最优解的极限为原问题的最优解.

例 4.8: 用外点罚函数法求解

$$\min\ f\left(\boldsymbol{x}\right) = \left(x_1 - 3\right)^2 + \left(x_2 - 2\right)^2$$
$$\text{s.t. } g(\boldsymbol{x}) = x_1 + x_2 - 4 \leq 0$$

例 4.8: 用外点罚函数法求解

$$\begin{aligned} \min\ & f\left(\boldsymbol{x}\right) = \left(x_1 - 3\right)^2 + \left(x_2 - 2\right)^2 \\ \text{s.t.}\ & g(\boldsymbol{x}) = x_1 + x_2 - 4 \leq 0 \end{aligned}$$

解: 作辅助函数

$$F(\boldsymbol{x}, \sigma) = \left(x_1 - 3\right)^2 + \left(x_2 - 2\right)^2 + \sigma\left[\max\left\{0, x_1 + x_2 - 4\right\}\right]^2$$

即:

$$F(\boldsymbol{x}, \sigma) = \begin{cases} \left(x_1 - 3\right)^2 + \left(x_2 - 2\right)^2, & x_1 + x_2 - 4 \leq 0 \\ \left(x_1 - 3\right)^2 + \left(x_2 - 2\right)^2 + \sigma\left(x_1 + x_2 - 4\right)^2, & x_1 + x_2 - 4 > 0 \end{cases}$$

因此:

$$\frac{\partial F}{\partial x_1} = \begin{cases} 2(x_1 - 3), & x_1 + x_2 \leq 4 \\ 2(x_1 - 3) + 2\sigma(x_1 + x_2 - 4), & x_1 + x_2 > 4 \end{cases}$$

惩罚函数法–外点法

同理:

$$\frac{\partial F}{\partial x_2} = \begin{cases} 2(x_2-2), & x_1+x_2 \le 4 \\ 2(x_2-2)+2\sigma(x_1+x_2-4), & x_1+x_2 > 4 \end{cases}$$

令: $\frac{\partial F}{\partial x_1} = \frac{\partial F}{\partial x_2} = 0$

得: $x_1(\sigma) = \frac{5\sigma+3}{2\sigma+1}$, $x_2(\sigma) = \frac{3\sigma+2}{2\sigma+1}$

又因该点处

$$\frac{\partial^2 F}{\partial x_1^2} = 2(\sigma+1), \quad \frac{\partial^2 F}{\partial x_2^2} = 2(\sigma+1), \quad \frac{\partial^2 F}{\partial x_1 \partial x_2} = 2\sigma$$

并且$\sigma > 0$, 可见$\nabla^2 F = \begin{pmatrix} 2(\sigma+1) & 2\sigma \\ 2\sigma & 2(\sigma+1) \end{pmatrix}$ 正定.

惩罚函数法–外点法

因此 $F(\boldsymbol{x},\sigma)$ 在点

$$\boldsymbol{x}(\sigma)=\left(\frac{5\sigma+3}{2\sigma+1},\frac{3\sigma+2}{2\sigma+1}\right)^T$$

处取得极小值. 令 $\sigma\to+\infty$ 得:

$$\boldsymbol{x}^*=\lim_{\sigma\to\infty}\boldsymbol{x}(\sigma)=\left(\frac{5}{2},\frac{3}{2}\right)^T.$$

所以原问题的最优解及最优值分别为:

$$\boldsymbol{x}^*=\left(\frac{5}{2},\frac{3}{2}\right)^T,\quad f(\boldsymbol{x}^*)=\frac{1}{2}.$$

惩罚函数法–外点法

3. 数值解法

实际问题中，上述以辅助函数$F(\boldsymbol{x},\sigma)$为目标函数的无约束极小化问题的解析解$\boldsymbol{x}(\sigma)$往往很难获得，σ的选取对问题的求解十分重要，因此通常取一个趋于无穷大的严格递增正数列$\{\sigma_k\}$，然后分别求解

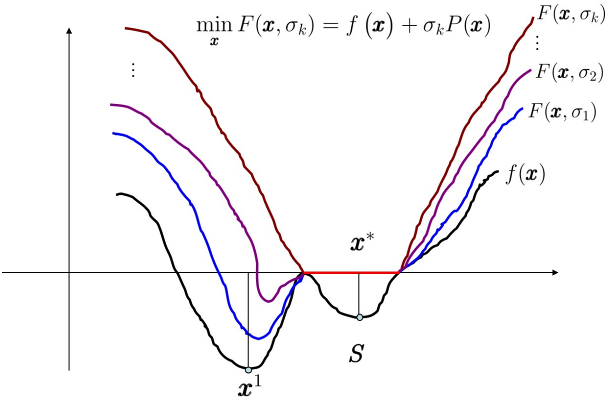
$$\min_{\boldsymbol{x}\in\mathbb{R}^n} F(\boldsymbol{x},\sigma_k) = f(\boldsymbol{x}) + \sigma_k P(\boldsymbol{x})$$

得极小点序列$\{\boldsymbol{x}^k\}$.

随着迭代次数k的增加，该点列渐渐收敛于原问题的极小点，由此产生SUMT外点法.

惩罚函数法–外点法

SUMT外点法:
改造曲线使得其在可行域S内不变, 在可行域S外将曲线逐步抬高.抬高到一定程度后, 此曲线的极小点就是原来函数在可行域S内的极小点.



惩罚函数法–外点法

以只具有等式约束的约束优化问题为例，从K-T条件角度分析和说明$\sigma \to +\infty$的原因：

原问题的KKT条件：

$$\nabla f(\boldsymbol{x}^*) + \sum_{j=1}^{p} v_j^* \nabla h_j(\boldsymbol{x}^*) = \mathbf{0}$$

$$h_j(\boldsymbol{x}^*) = 0$$

添加罚函数项无约束优化问题的KKT条件：

$$\nabla f(\boldsymbol{x}) + \sum_{j=1}^{p} \sigma h_j(x) \nabla h_j(\boldsymbol{x}) = \mathbf{0}$$

假设两个问题收敛到同一点，对比K-T条件(梯度式)，

则$\sigma h_j(\boldsymbol{x}) \approx v_j^*$.由于最优点处乘子$v_j^*$固定，为使约束$h_j(\boldsymbol{x}) = 0$成立，故$\sigma \to +\infty$.

惩罚函数法–外点法

以下继续分析和说明为什么应用惩罚函数法时会遭遇数值困难：
考虑辅助函数$F(\boldsymbol{x},\sigma)$的Hessian矩阵：

$$\nabla^2 F(\boldsymbol{x},\sigma) = \nabla^2 f(\boldsymbol{x}) + \sum_{i=1}^{p} \sigma h_j(\boldsymbol{x})\nabla^2 h_j(\boldsymbol{x}) + \sigma \nabla h_j(\boldsymbol{x})\nabla h_j(\boldsymbol{x})^T.$$

等号右边的前两项可以使用拉格朗日函数$\mathcal{L}(\boldsymbol{x},\boldsymbol{v}^*)$来近似，即：

$$\nabla^2 F(\boldsymbol{x},\sigma) \approx \nabla^2 \mathcal{L}\left(\boldsymbol{x},\boldsymbol{v}^*\right) + \sigma \nabla h_j(\boldsymbol{x})\nabla h_j(\boldsymbol{x})^T.$$

可见，右端矩阵为一个定值矩阵和一个最大特征值趋于正无穷的矩阵 之和.这导致$\nabla^2 F(\boldsymbol{x},\sigma)$的条件数无限增大，不管利用何种无约束优化算法，求解子问题的难度均大大增加.因此在实际应用中，我们不能令惩罚因子σ趋于正无穷.

惩罚函数法–外点法

SUMT外点法(数值解法)

1) 初始化$\boldsymbol{x}^0$, 初始惩罚因子$\sigma_1 > 0$, 放大因子$C > 1$, 允许误差$\epsilon > 0$, 令$k = 1$.
2) 以$\boldsymbol{x}^{k-1}$为初始点, 用解析法求驻点或利用无约束优化方法求解
$$\min_{\boldsymbol{x}\in\mathbb{R}^n} F(\boldsymbol{x}, \sigma_k) = f(\boldsymbol{x}) + \sigma_k P(\boldsymbol{x})$$
得其极小点$\boldsymbol{x}(\sigma_k)$并记为$\boldsymbol{x}^k$.
3) 若$\sigma_k P(\boldsymbol{x}^k) < \epsilon$, 停止运算, 取$\boldsymbol{x}^* = \boldsymbol{x}^k$作为近似解.
否则, 令$\sigma_{k+1} = C\sigma_k$, $k = k + 1$, 返回Step 2).

【注】根据经验, 常常取$\sigma_1 = 1$, $C \in [5, 10]$.

惩罚函数法–外点法

引理 4.1:
对于SUMT外点法产生的点列$\{\boldsymbol{x}^k\}$及$\sigma_{k+1} > \sigma_k$，下列不等式成立:

(1) $F(\boldsymbol{x}^k, \sigma_k) \le F(\boldsymbol{x}^{k+1}, \sigma_{k+1})$ $\leftrightarrow$整体目标函数值上升;

(2) $P(\boldsymbol{x}^k) \ge P(\boldsymbol{x}^{k+1})$ $\leftrightarrow$迭代点列逐渐收敛到可行域S边界;

(3) $f(\boldsymbol{x}^k) \le f(\boldsymbol{x}^{k+1})$ $\leftrightarrow$原始目标函数逐渐增大.

证: (1)

$$\begin{aligned} F(\boldsymbol{x}^{k+1}, \sigma_{k+1}) &= f(\boldsymbol{x}^{k+1}) + \sigma_{k+1} P(\boldsymbol{x}^{k+1}) \\ &\ge f(\boldsymbol{x}^{k+1}) + \sigma_k P(\boldsymbol{x}^{k+1}) \\ &\ge f(\boldsymbol{x}^k) + \sigma_k P(\boldsymbol{x}^k) = F(\boldsymbol{x}^k, \sigma_k). \end{aligned}$$

惩罚函数法–外点法

引理 4.1:
对于SUMT外点法产生的点列$\{\boldsymbol{x}^k\}$及$\sigma_{k+1} > \sigma_k$，下列不等式成立：

(1) $F(\boldsymbol{x}^k, \sigma_k) \leq F(\boldsymbol{x}^{k+1}, \sigma_{k+1}) \leftrightarrow$整体目标函数值上升;
(2) $P(\boldsymbol{x}^k) \geq P(\boldsymbol{x}^{k+1}) \leftrightarrow$迭代点列逐渐收敛到可行域$S$边界;
(3) $f(\boldsymbol{x}^k) \leq f(\boldsymbol{x}^{k+1}) \leftrightarrow$原始目标函数逐渐增大.

证： (2)
$$f(\boldsymbol{x}^{k+1}) + \sigma_{k+1} P(\boldsymbol{x}^{k+1}) \leq f(\boldsymbol{x}^k) + \sigma_{k+1} P(\boldsymbol{x}^k),$$
$$f(\boldsymbol{x}^k) + \sigma_k P(\boldsymbol{x}^k) \leq f(\boldsymbol{x}^{k+1}) + \sigma_k P(\boldsymbol{x}^{k+1}).$$

将上述两个式子相加并合并同类项，可得
$$\sigma_{k+1} P(\boldsymbol{x}^{k+1}) + \sigma_k P(\boldsymbol{x}^k) \leq \sigma_{k+1} P(\boldsymbol{x}^k) + \sigma_k P(\boldsymbol{x}^{k+1}).$$

即$(\sigma_{k+1} - \sigma_k) P(\boldsymbol{x}^{k+1}) \leq (\sigma_{k+1} - \sigma_k) P(\boldsymbol{x}^k) \leftrightarrow P(\boldsymbol{x}^{k+1}) \leq P(\boldsymbol{x}^k)$.

惩罚函数法–外点法

引理 4.1:
对于SUMT外点法产生的点列$\{\boldsymbol{x}^k\}$及$\sigma_{k+1} > \sigma_k$, 下列不等式成立:

(1) $F(\boldsymbol{x}^k, \sigma_k) \leq F(\boldsymbol{x}^{k+1}, \sigma_{k+1})$$\leftrightarrow$整体目标函数值上升;
(2) $P(\boldsymbol{x}^k) \geq P(\boldsymbol{x}^{k+1})$$\leftrightarrow$迭代点列逐渐收敛到可行域$\boldsymbol{S}$边界;
(3) $f(\boldsymbol{x}^k) \leq f(\boldsymbol{x}^{k+1})$$\leftrightarrow$原始目标函数逐渐增大.

证: (3)

$$f(\boldsymbol{x}^{k+1}) + \sigma_k P(\boldsymbol{x}^{k+1}) \geq f(\boldsymbol{x}^k) + \sigma_k P(\boldsymbol{x}^k),$$
$$\sigma_k P(\boldsymbol{x}^{k+1}) \leq \sigma_k P(\boldsymbol{x}^k).$$

综合上述两个式子, 可得

$$f(\boldsymbol{x}^{k+1}) \geq f(\boldsymbol{x}^k).$$

惩罚函数法–外点法

引理 4.2：
设$\boldsymbol{x}^*$是约束优化问题(I)的一个最优解，则对于SUMT外点法产生的点列$\{\boldsymbol{x}^k\}$及每一个k，有下列不等式成立：

$$f(\boldsymbol{x}^*) \geq F(\boldsymbol{x}^k, \sigma_k) \geq f(\boldsymbol{x}^{k+1}).$$

证： 因为$\boldsymbol{x}^*$是约束优化问题(I)的一个最优解，所以$\boldsymbol{x}^* \in \boldsymbol{S}$，$P(\boldsymbol{x}^*) = 0$. 因此

$$\begin{aligned} f(\boldsymbol{x}^*) &= f(\boldsymbol{x}^*) + \sigma_k P(\boldsymbol{x}^*) \\ &\geq f(\boldsymbol{x}^k) + \sigma_k P(\boldsymbol{x}^k) \\ &= F(\boldsymbol{x}^k, \sigma_k) \\ &\geq f(\boldsymbol{x}^k). \end{aligned}$$

惩罚函数法–外点法

定理 4.5：收敛性定理

设$\min\limits_{\boldsymbol{x}\in\mathbb{R}^n} F(\boldsymbol{x},\sigma_k)$的极小点为$\boldsymbol{x}^k=\boldsymbol{x}(\sigma_k)$, 则外点法迭代过程中会出现以下两种情况:

(1) 如果存在某个$k_0\geq 1$满足$\boldsymbol{x}^{k_0}\in\boldsymbol{S}$, 则$\boldsymbol{x}^*=\boldsymbol{x}^{k_0}$;

(2) 如果上述情况不发生, 则得到一个无穷点列$\{\boldsymbol{x}^k\}_{k=1}^{+\infty}$.当$f(\boldsymbol{x})$, $g_i(\boldsymbol{x})$, $h_j(\boldsymbol{x})$是定义在可行域$\boldsymbol{S}$上的连续函数时, 有$\boldsymbol{x}^*=\lim\limits_{k\to+\infty}\boldsymbol{x}^k$.

惩罚函数法–外点法

优点

- 初始点可以任选;
- 对等式约束和不等式约束均可适用;
- 若有了求解无约束问题的好算法, 则利用外罚函数法求解约束问题很方便.

惩罚函数法–外点法

缺点

- 外点法不适用于可行域外$f(\boldsymbol{x})$的性质复杂或者没有定义的情况;
- 每个近似解$\boldsymbol{x}^k$可能不是可行解，这是某些实际问题所无法接受的(内点罚函数法可以解决);
- 初始罚因子σ_1要选择得当，否则会影响迭代计算的正常进行.若σ_1太小，将增加迭代的次数; 若σ_1太大，会使惩罚函数的性态变坏，甚至难以收敛到极值点; 许多计算表明，取$\sigma_1 = 1$常常可以取得满意的效果;
- 理论上可证明σ_k取越大越好，而σ_k越大将导致目标函数$F(\boldsymbol{x}, \sigma)$的 Hessian矩阵条件数越大、趋于病态，给无约束问题求解增加很大困难，甚至无法求解(乘子法法可以解决).

惩罚函数法–内点法

二、内点罚函数法(碰壁函数法)–内点法

1. 内点法基本思想

在可行域的边界筑起一道很高的“围墙”，当迭代点靠近可行域的边界时, 目标函数值陡然增大, 这相当于对它进行惩罚, 从而阻止迭代点穿越边界, 这样就可以把最优解“挡”在可行域内.

【注】内点法要求迭代点在可行域内部移动.这要求初始点必须是内点, 并且可行域的内部必须是非空的.而等式约束构成的集合内部为空.因此, 内点法只适合于不等式约束问题.另一方面, 可行域的内点集S中的点可以任意接近于S的任一点, 即S不能包含孤立点或孤立线段.

惩罚函数法–内点法

从数学角度而言，内点法从一个可行点$\boldsymbol{x}^0 \in \boldsymbol{S}$出发，通过求辅助函数极小点得新的可行点，即在可行点之间进行迭代，最终使迭代点逐渐靠近或收敛于最优点.内点罚函数法又称为障碍函数法.
针对

$$\min_{\boldsymbol{x}} \; f(\boldsymbol{x}) \; s.t. \; g_i(\boldsymbol{x}) \leq 0, \; i = 1, 2, \cdots m$$

作辅助函数$F(\boldsymbol{x}, r) = f(\boldsymbol{x}) + rB(\boldsymbol{x})$，$r > 0$为碰壁/障碍因子.其中，惩罚函数$B(\boldsymbol{x})$满足：

1. $B(\boldsymbol{x})$在可行域$\boldsymbol{S}$内部是连续的;
2. 对可行域$\boldsymbol{S}$内部的$\boldsymbol{x}$，$B(\boldsymbol{x})$为幅值很小的数;
3. 当$\boldsymbol{x}$趋于$\boldsymbol{S}$的边界时，$B(\boldsymbol{x})$迅速趋于无穷大.

惩罚函数法–内点法

构造辅助函数：

$$F(\boldsymbol{x},r) = f(\boldsymbol{x}) + rB(\boldsymbol{x})$$

其中$r>0$为很小的数，

$$B(\boldsymbol{x})=\sum_{i=1}^{m}-\frac{1}{g_i(\boldsymbol{x})}, \qquad \boxed{\text{倒数障碍函数}}$$

或

$$B(\boldsymbol{x})=-\sum_{i=1}^{m}\ln(-g_i(\boldsymbol{x})). \boxed{\text{对数障碍函数}}$$

惩罚函数法–内点法

辅助函数：

$$F(\boldsymbol{x}, r) = f(\boldsymbol{x}) + rB(\boldsymbol{x})$$

其中$r > 0$为很小的数.

分析：

- 当$\boldsymbol{x}$为可行域$\boldsymbol{S}$的内点且远离可行域边界时, $rB(\boldsymbol{x})$为幅值很小的数, 几乎不受惩罚.此时, $F(\boldsymbol{x}, r)$的取值近似原目标函数;
- 当$\boldsymbol{x}$接近边界时, $B(\boldsymbol{x}) \to +\infty$, 施以很重的惩罚$rB(\boldsymbol{x})$.

因此, 按上策略对从内部企图穿越可行域边界的点在目标函数中加入障碍, 距边界越近, 障碍越大, 在边界上给予无穷大的障碍.从而保证迭代点一直在可行域内部移动, 故称此类算法为内点法.

2. 解析法

$$\min_{\boldsymbol{x}} \ f(\boldsymbol{x}) \ s.t. \ g_i(\boldsymbol{x}) \leq 0, \ i = 1, 2, \cdots, m$$

(1) 构造辅助函数：$F(\boldsymbol{x}, r) = f(x) + rB(\boldsymbol{x})$,其中$r > 0$为很小的正数; $B(\boldsymbol{x}) = \sum_{i=1}^{m} -\frac{1}{g_i(\boldsymbol{x})}$或$B(\boldsymbol{x}) = -\sum_{i=1}^{m} \ln(-g_i(\boldsymbol{x}))$.

(2) 求解：$\min_{\boldsymbol{x} \in \boldsymbol{S}} F(\boldsymbol{x}, r) = f(\boldsymbol{x}) + rB(\boldsymbol{x})$得最优解$\boldsymbol{x}(r)$.

(3) 令$r \to 0$, 有$\boldsymbol{x}(r) \to \boldsymbol{x}^*$, 即得原问题的(近似)最优解.

【注】从形式上看，$\min_{\boldsymbol{x} \in \boldsymbol{S}} F(\boldsymbol{x}, r)$仍是一个约束优化问题; 从计算的观点看, 该问题是一个无约束优化问题. 原因是, 在可行域的边界附近，目标函数值$F(\boldsymbol{x}, r)$趋于$+\infty$, 只要从可行域$\boldsymbol{S}$的任何一个内点开始迭代, 并注意控制一维搜索的步长, 就可使得迭代点不越过可行域.

例 4.9: 用内点罚函数法求解:

$$
\begin{aligned}
\min\ & f(\boldsymbol{x}) = \frac{1}{3}(x_1+1)^3 + x_2 \\
\text{s.t. } & 1 - x_1 \le 0, \\
& x_2 \ge 0
\end{aligned}
$$

惩罚函数法–内点法

例 4.9: 用内点罚函数法求解:

$$
\begin{aligned}
\min\ f(\boldsymbol{x}) &= \frac{1}{3}(x_1+1)^3 + x_2 \\
\text{s.t. } 1 - x_1 &\le 0, \\
x_2 &\ge 0
\end{aligned}
$$

解: 作辅助函数

$$
F(\boldsymbol{x}, r) = \frac{1}{3}(x_1+1)^3 + x_2 + r\left(\frac{1}{x_1-1} + \frac{1}{x_2}\right)
$$

其中 $r > 0$ 为很小的正数.

令: $\frac{\partial F}{\partial x_1} = (x_1+1)^2 - \frac{r}{(x_1-1)^2} = 0$, $\frac{\partial F}{\partial x_2} = 1 - \frac{r}{x_2^2} = 0$

惩罚函数法–内点法

再注意到 $\boldsymbol{x} \in \boldsymbol{S}$, 则 $x_1 > 1$.

所以 $\boldsymbol{x}(r) = \begin{pmatrix} \sqrt{1+\sqrt{r}} \\ \sqrt{r} \end{pmatrix}$

可验证该点处的Hessian矩阵

$$\nabla^2 F = \begin{pmatrix} 2\left(\sqrt{1+\sqrt{r}}+1\right) + \frac{2r}{\left(\sqrt{1+\sqrt{r}}-1\right)^3} & 0 \\ 0 & \frac{2}{\sqrt{r}} \end{pmatrix}$$

正定, 所以 $\boldsymbol{x}(r)$ 是 $F(\boldsymbol{x}, r)$ 的极小点. 令 $r \to 0$ 有:
$\boldsymbol{x}(r) \to \boldsymbol{x}^* = (1,0)^T$, $f(\boldsymbol{x}^*) = \frac{8}{3}$

惩罚函数法–内点法

例 4.10: 用内点罚函数法求解:

$$\min\ f(\boldsymbol{x}) = x_1^2 + x_2^2$$
$$s.t\ x_1 \geq 1$$

解: 用 $B(\boldsymbol{x}) = -\sum_{i=1}^{m} \ln(-g_i(\boldsymbol{x}))$ 进行求解.

惩罚函数法—内点法

3. 数值算法

对于内点法, r越小, 所构造的无约束问题的最优解越接近原约束优化的最优解.因此, r的选取对问题的求解十分重要.一般做法是取一个收敛于零的严格递减正数列$\{r_k\}$:

$$r_k > 0\text{且}r_1 > r_2 > r_3 > \cdots > r_k \to 0.$$

然后依次求解

$$\min_{\boldsymbol{x}\in\boldsymbol{S}} F(\boldsymbol{x}, r_k) = f(\boldsymbol{x}) + r_k P(\boldsymbol{x})$$

得极小点序列$\{\boldsymbol{x}^k\}$.随着迭代次数k的增加, 该点列渐渐收敛于原问题的极小点.由此产生如下SUMT内点法(数值解法).

惩罚函数法—内点法

SUMT内点法(数值解法)

1) 初始化$\boldsymbol{x}^0 \in \mathbb{R}^n$, $\epsilon > 0$, 障碍因子r_1, 缩小系数c, 令$k = 1$.
2) 以$\boldsymbol{x}^{k-1}$为初始点, 用解析法求驻点或利用无约束优化方法求解

$$\min_{\boldsymbol{x} \in \boldsymbol{S}} F(\boldsymbol{x}, r) = f(\boldsymbol{x}) + r_k B(\boldsymbol{x})$$

得其最优解$\boldsymbol{x}(r_k)$并记为$\boldsymbol{x}^k$.
3) 若$\left| r_k B\left(\boldsymbol{x}^k\right) \right| < \epsilon$, 则$\boldsymbol{x}^* = \boldsymbol{x}^k$, 停止迭代.
否则, 令$r_{k+1} = c r_k$, $k = k + 1$, 转Step 2).

【注】根据经验, 常常取$r_1 = 10$, $c = 0.1$.

惩罚函数法–内点法

引理 4.3:
对于SUMT内点法产生的点列$\{\boldsymbol{x}^k\}$, 及$r_{k+1} > r_k$, 下列不等式成立:

(1) $F(\boldsymbol{x}^{k+1}, r_{k+1}) \le F(\boldsymbol{x}^k, r_k)$$\leftrightarrow$整体目标函数值下降;
(2) $B(\boldsymbol{x}^{k+1}) \ge B(\boldsymbol{x}^k)$$\leftrightarrow$迭代点列逐渐收敛到可行域$S$边界;
(3) $f(\boldsymbol{x}^{k+1}) \le f(\boldsymbol{x}^k)$$\leftrightarrow$原始目标函数逐渐变小.

惩罚函数法–内点法

引理 4.4:

设$\boldsymbol{x}^*$是约束优化问题$\min\ f(\boldsymbol{x})\ s.t.\ g_i(\boldsymbol{x}) \le 0,\ i=1,2,\cdots,m$的一个最优解, 则对于SUMT内点法产生的点列$\{\boldsymbol{x}^k\}$及每一个$k$, 有下列不等式成立:

$$f(\boldsymbol{x}^*) \le f(\boldsymbol{x}^k) \le F(\boldsymbol{x}^k, r_k).$$

惩罚函数法–内点法

定理 4.6：收敛性定理

若$f(\boldsymbol{x})$, $g_i(\boldsymbol{x})$为定义在可行域S上的连续函数.对于任意初始可行点, 记由SUMT内点罚函数法产生的点列为$\{\boldsymbol{x}^k\}$, 则

$$\boldsymbol{x}^* = \lim_{k\to+\infty} \boldsymbol{x}^k.$$

惩罚函数法–内点法

注意事项

- 使用内点法时, 初始点x^0应选择一个离约束边界较远的可行点.如太靠近某一约束边界, 构造的惩罚函数可能由于障碍项的值很大而变得畸形, 使求解无约束优化问题发生困难;
- 应适当设定障碍因子的初值r_1, 否则会影响迭代计算的正常进行.一般而言, 太大将增加迭代次数; 太小会使惩罚函数的性态变坏, 甚至难以收敛到极值点无一般性的有效方法.对于不同的问题, 都要经过多次试算, 才能决定一个适当的r_1.

惩罚函数法–内点法

- 惩罚因子的缩减系数c的选取
  在构造序列惩罚函数时, 惩罚因子r逐渐递减到0, 相邻两次迭代的惩罚因子的关系为:
  $$r^k = cr^{k-1} \quad (k = 1, 2, ...)$$
  其中, c为缩减系数.一般而言, c值的大小在迭代过程中不起决定性作用, 其取值通常小于1且满足$c \in [0.1, 0.7]$;
- 内点法适于解仅含不等式约束问题, 并且每次迭代的点都是可行点.但要求初始点为可行域的内点, 需要浪费相当的工作量.

惩罚函数法–对比

外点法	内点法
初始点可以任选	初始点必须在可行域的内部
对具有等式约束和不等式约束的数学模型均适用	仅适用于具有不等式约束的数学模型
仅最优解为可行设计方案	迭代过程中各个点均为可行设计方案
一般收敛较快	一般收敛较慢
迭代中x^k不在可行域**S**中	迭代中x^k在可行域**S**中
罚因子为递增，递增率**c**，有**c>1**	罚因子为递减，递减率**c**，有**0<c<1**
非凸规划适用	非凸规划适用

惩罚函数法–混合惩罚函数法

三、混合惩罚函数法

基本思想：把外点罚函数与内点罚函数结合起来，构造新的罚函数. 设当前迭代点为$\boldsymbol{x}^{k-1}$，对于$\boldsymbol{x}^{k-1}$满足的那些不等式约束，用内点法来构造碰壁项$B(\boldsymbol{x})$，对于$\boldsymbol{x}^{k-1}$不满足的不等式约束和等式约束，按外点法构造惩罚项$P(\boldsymbol{x})$，即定义混合罚函数及辅助函数为：

$$F(\boldsymbol{x}, r_k) = f(\boldsymbol{x}) + r_k B(\boldsymbol{x}) + \frac{1}{r_k} P(\boldsymbol{x}).$$

惩罚函数法–混合惩罚函数法

对于新的辅助函数

$$F\left(\boldsymbol{x}, r_k\right)=f\left(\boldsymbol{x}\right)+r_k B\left(\boldsymbol{x}\right)+\frac{1}{r_k} P\left(\boldsymbol{x}\right)$$

其中，$r_k>0$为递减数列，且$\lim_{k\rightarrow+\infty} r_k=0$;

$$B\left(\boldsymbol{x}\right)=\sum_{i\in\boldsymbol{I}_1}-\frac{1}{g_i\left(\boldsymbol{x}\right)},\ \boldsymbol{I}_1=\left\{i\left|g_i(\boldsymbol{x}^{k-1})<0\right.\right\};$$

$$P(\boldsymbol{x})=\sum_{i\in\boldsymbol{I}_2}\left[\max\left\{0, g_i\left(\boldsymbol{x}\right)\right\}\right]^2+\sum_{j=1}^{p} h_j^2\left(\boldsymbol{x}\right),\ \boldsymbol{I}_2=\left\{i\left|g_i(\boldsymbol{x}^{k-1})\geq 0\right.\right\};$$

$$\boldsymbol{I}_1\bigcup\boldsymbol{I}_2=\{1,2,\cdots,m\}.$$

惩罚函数法–混合惩罚函数法

SUMT混合惩罚函数法(数值解法)

1) 初始化$\boldsymbol{x}^0 \in \mathbb{R}^n$, 初始惩罚因子$r_0 > 0$, 允许误差$\epsilon_1, \epsilon_2 > 0$, 缩小因子$c(0 < c < 1)$, 令 $k = 1$.
2) 以$\boldsymbol{x}^{k-1}$为初始点, 用解析法求驻点或利用无约束优化方法求解:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} F\left(\boldsymbol{x}, r_k\right) = f\left(\boldsymbol{x}\right) + r_k B\left(\boldsymbol{x}\right) + \frac{1}{r_k} P(\boldsymbol{x}),$$

   得最优解$\boldsymbol{x}(r_k)$并记为$\boldsymbol{x}^k$.
3) 若$\left|r_k B\left(\boldsymbol{x}^k\right)\right| < \epsilon_1, \frac{1}{r_k} P\left(\boldsymbol{x}^k\right) < \epsilon_2$, 则$\boldsymbol{x}^* = \boldsymbol{x}^k$, 停止迭代; 否则,令$r_{k+1} = c r_k, k = k + 1$, 转Step 2).